Data Quality Report

1. Summary of the dataset

Dataset *card transactions* contains 10 variables and 96,708 observations. The 10 variables are **'Recordnum', 'Cardnum', 'Date', 'Merchantnum', 'Merch Description', 'Merchant State', 'Merchant Zip', 'Transtype', 'Amount', 'Fraud'.**

Below are summaries of numeric variables and non-numeric variables.

Summary of Numeric Variable

	count	mean	std	min	50%	max	populated_rate
Amount	96708.00	427.87	10008.47	0.01	137.90	3102045.53	1.00

Summary of Non-Numeric Variables

	count	unique	populated_rate
Recordnum	96708	96708	1.00
Cardnum	96708	1644	1.00
Date	96708	365	1.00
Merchantnum	93333	13090	0.97
Merch Description	96708	13125	1.00
Merchant State	95513	227	0.99
Merchant Zip	96708	4568	1.00
Transtype	96708	4	1.00
Fraud	96708	2	1.00

2. Detailed information for Each Field

Field Name	*Description*
Recordnum	Index of the dataset

Field Name	Description
Cardnum	Credit card number of the transaction

Top 20 Cardnum

Cardnum	Counts
5142131548	
5142132941	
5142176897	
5142273878	
5142259479	
5142125684	
5142143744	
5142152067	
5142157720	
5142230181	
5142190147	
5142149691	
5142189945	
5142299634	
5142187452	
5142223373	
5142297710	
5142189108	
5142184598	
5142148452	
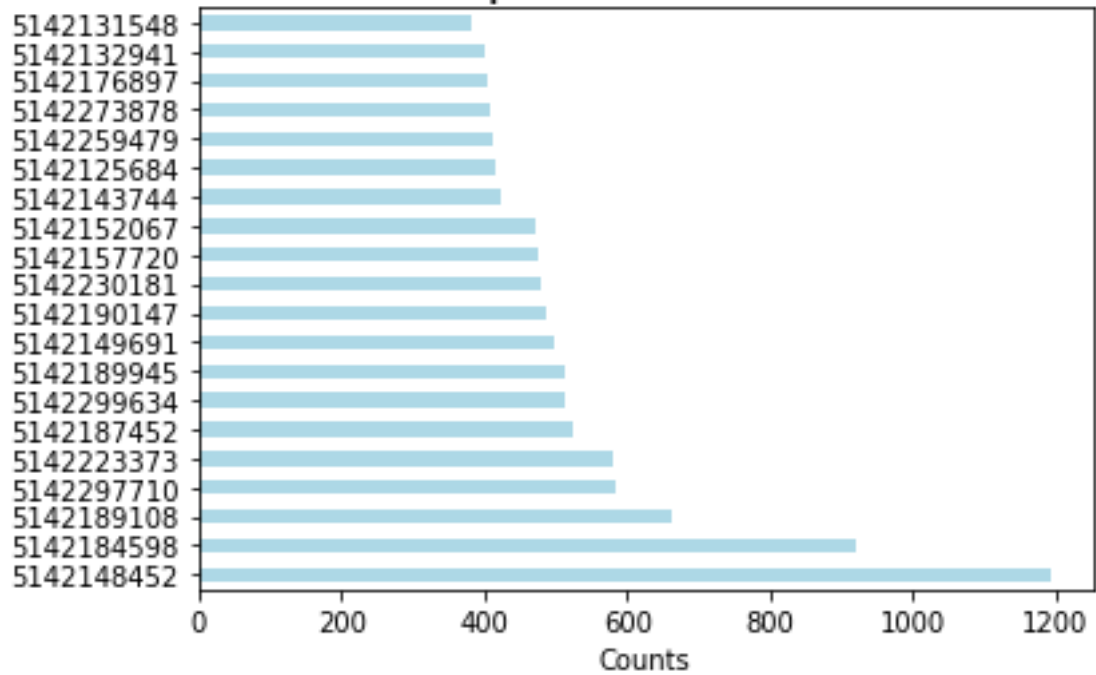
Field Name	Description
Date	Dates of transactions

Distribution of Transaction Dates
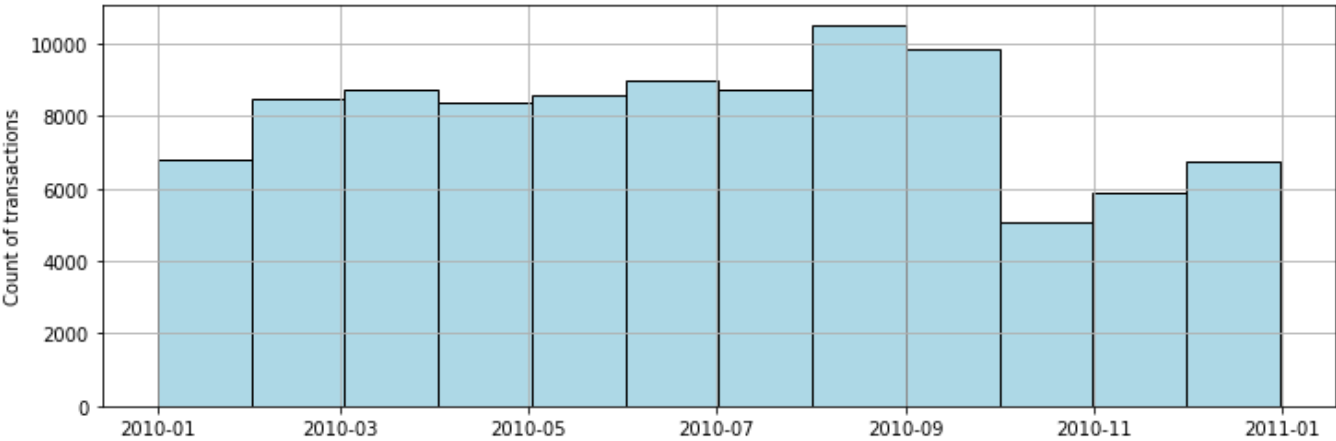
Field Name	Description
Merchantnum	Unique numbers of transaction merchants

Top 20 Merchantnum
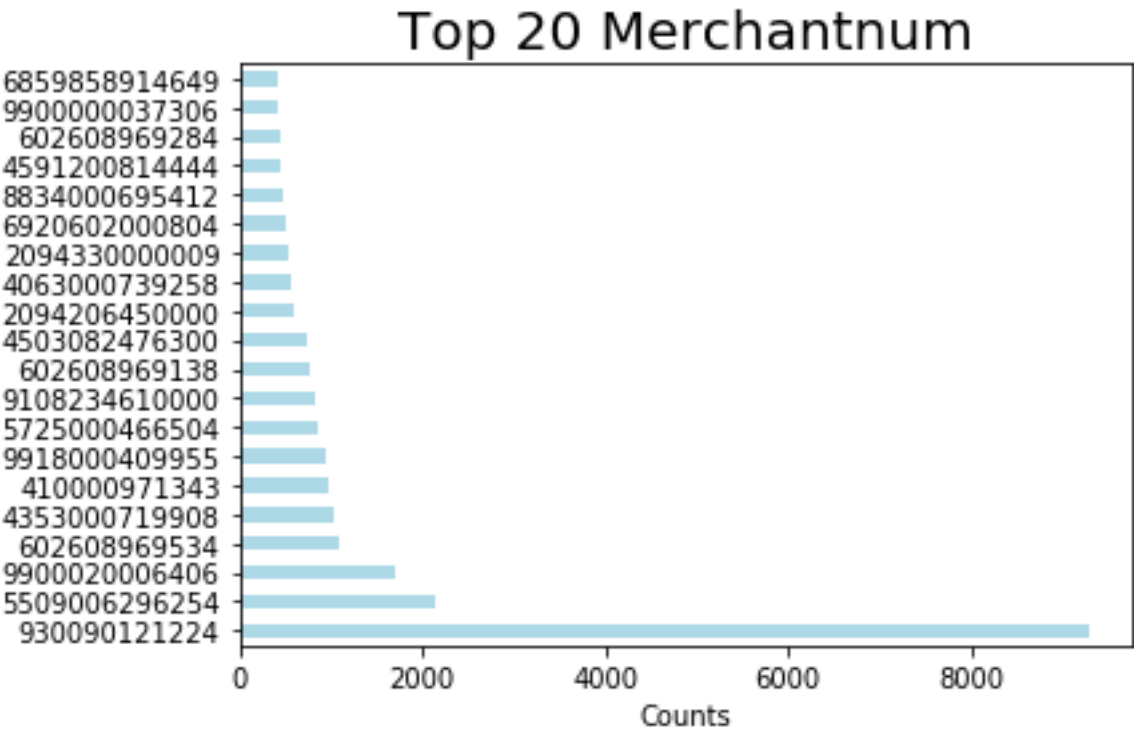
Field Name	Description
Merch Description	Descriptions of transaction merchants

Top 20 Merch Description
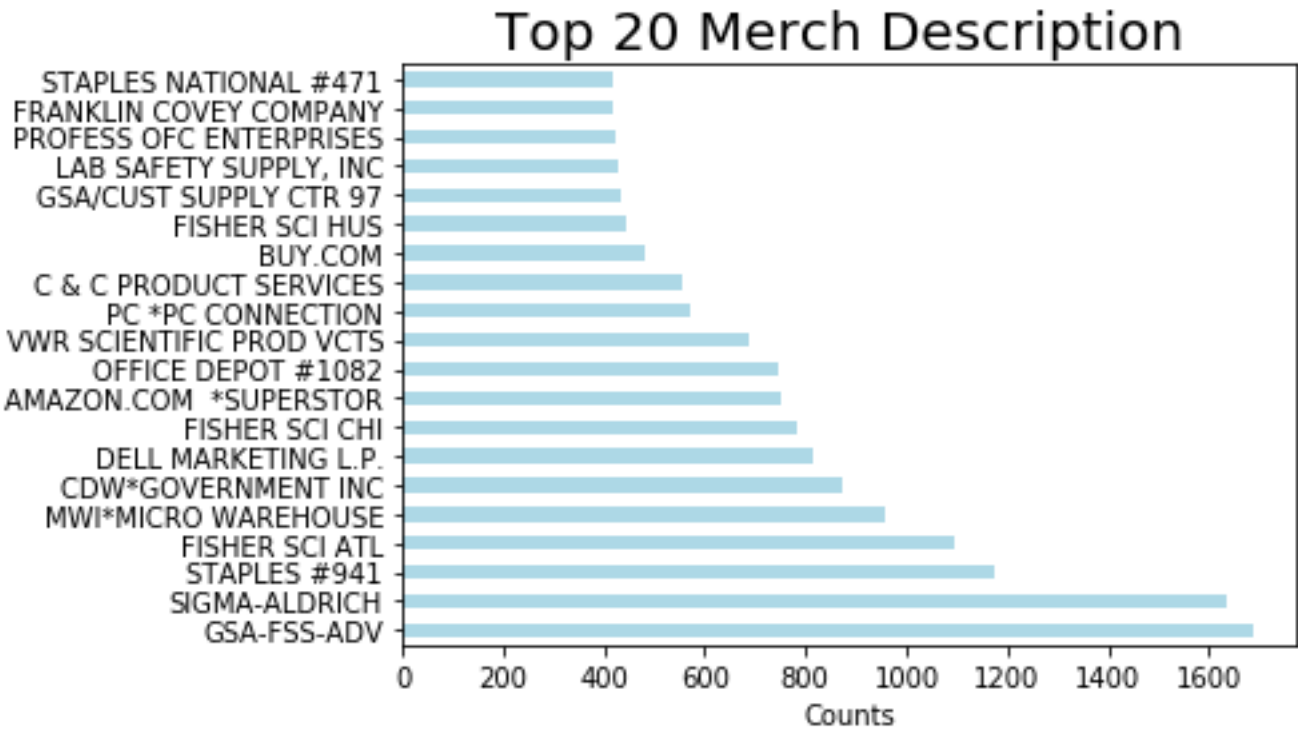
Field Name	Description
Merchant State	State of transaction merchants

Top 20 Merchant State
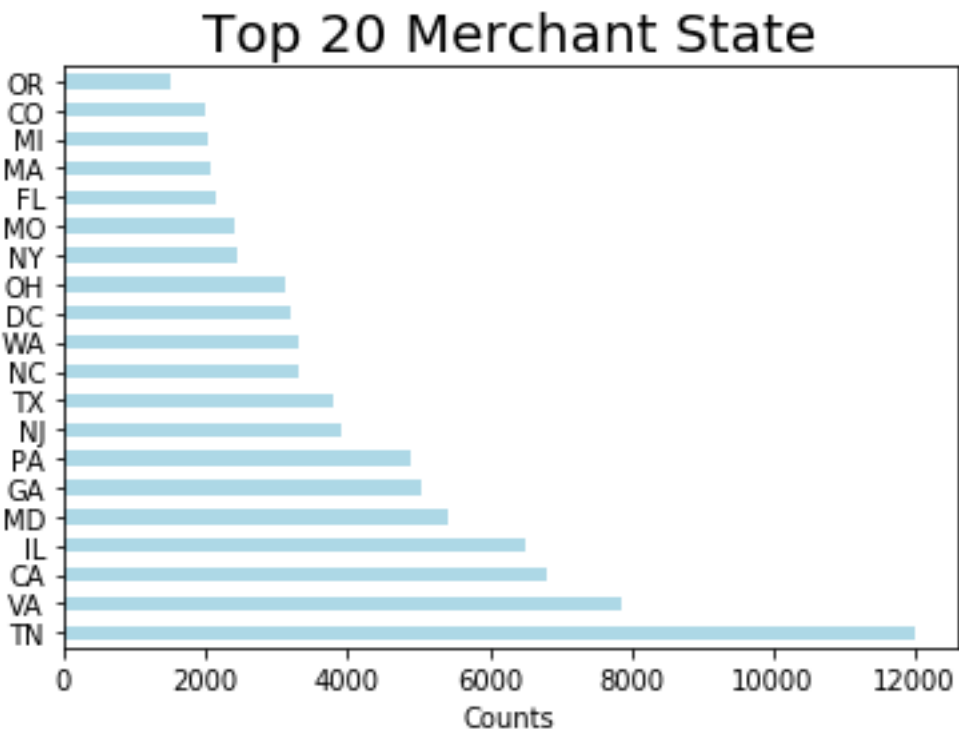
Field Name	Description
Merchant Zip	Zip codes of transaction merchants

Top 20 Merchant Zip
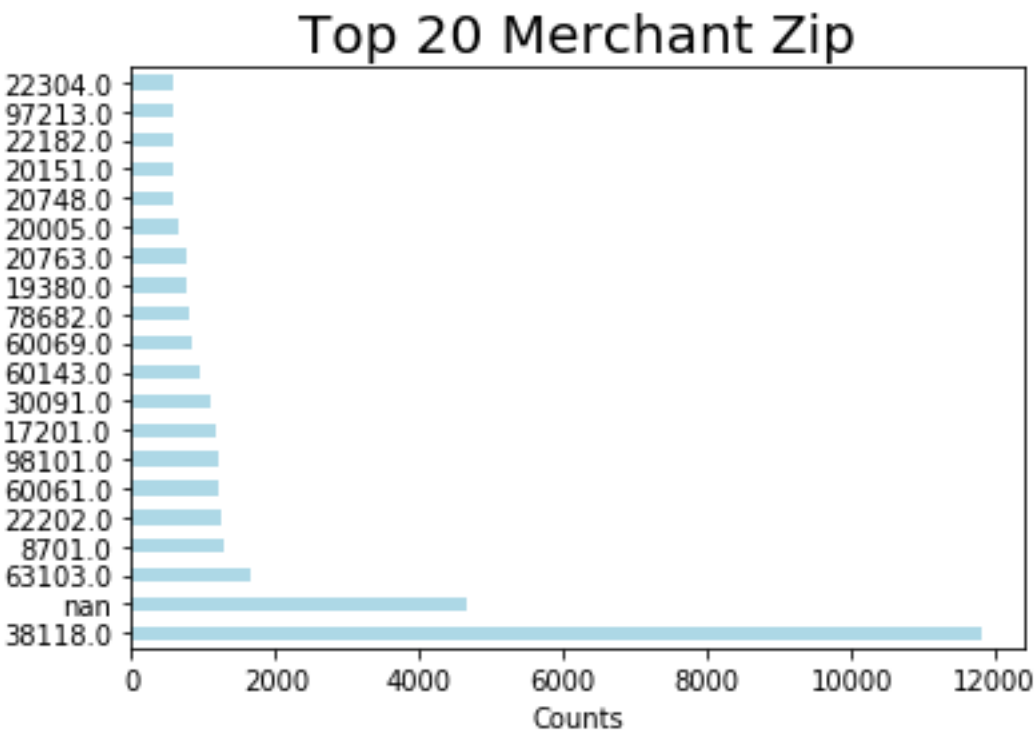
Field Name	Description
Transtype	Types of transactions

Transtype counts
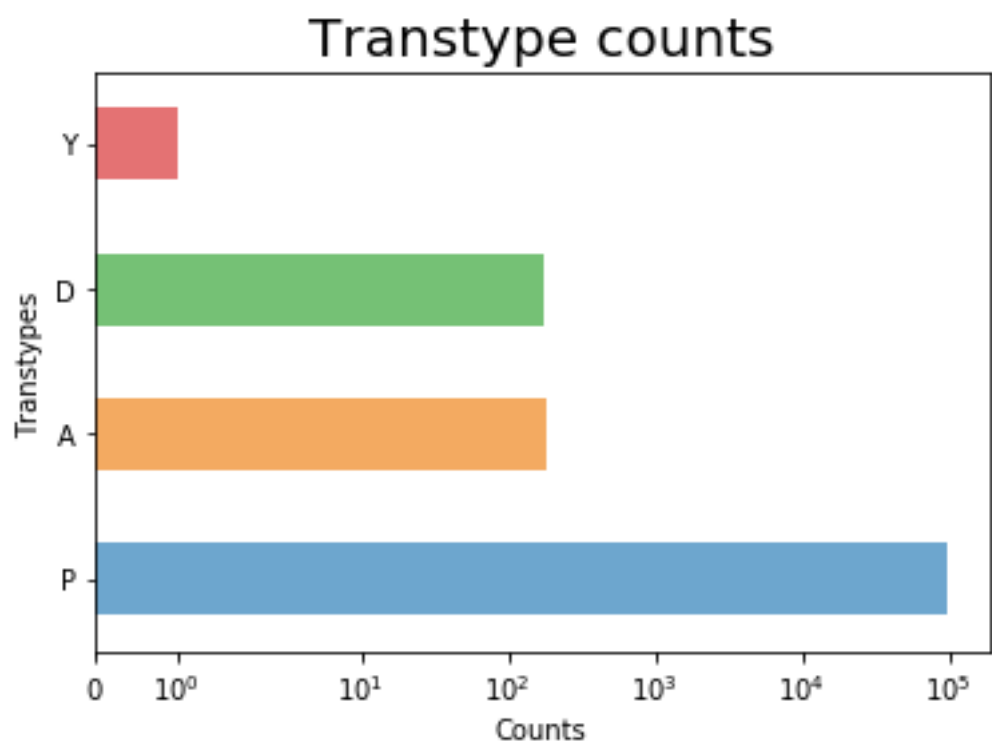
Field Name	Description
Amount	Amount of transactions

Transaction Amount Distribution
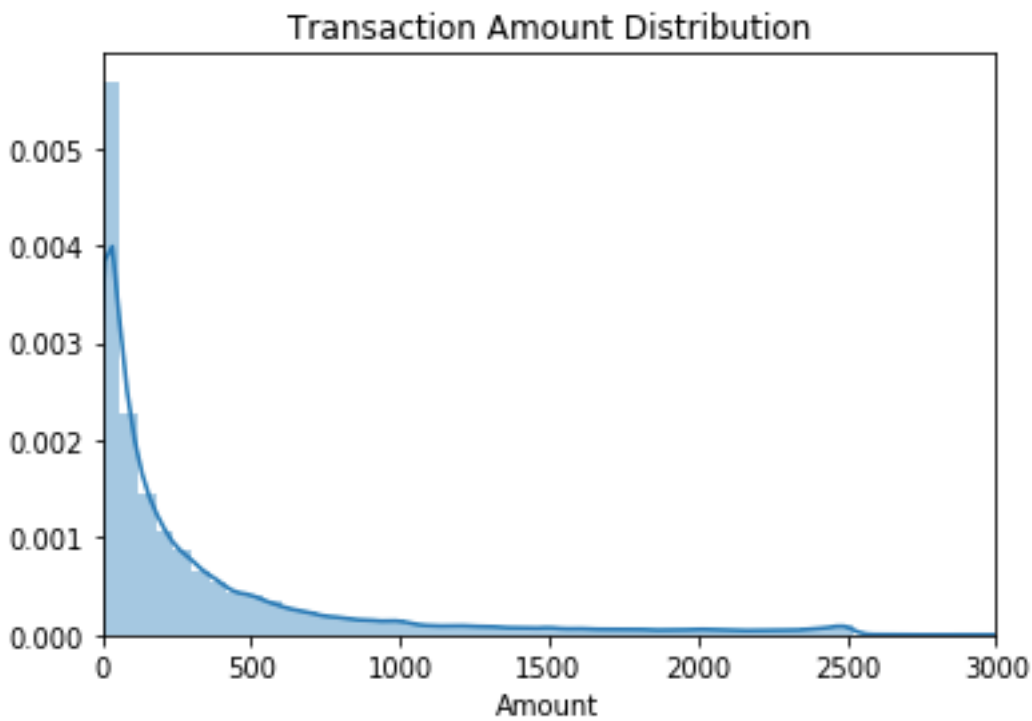
NOTE
Record **52593** is an outlier in **Amount**. The reason of its high **Amount** value is because the transaction was executed with Mexico Pesos.

	Top 10 Amount			52593
0	3102045.53		Recordnum	52594
1	47900.00		Cardnum	5142189135
2	30372.46		Date	2010-07-13 00:00:00
3	28392.84		Merchantnum	NaN
4	27218.00		Merch Description	INTERMEXICO
5	26910.00		Merchant State	NaN
6	25805.35		Merchant Zip	nan
7	25000.00		Transtype	P
8	24990.00		Amount	3102045.53
9	24946.90		Fraud	0

Field Name	Description
Fraud	Whether a certain record is a fraud or not.

Fraud counts